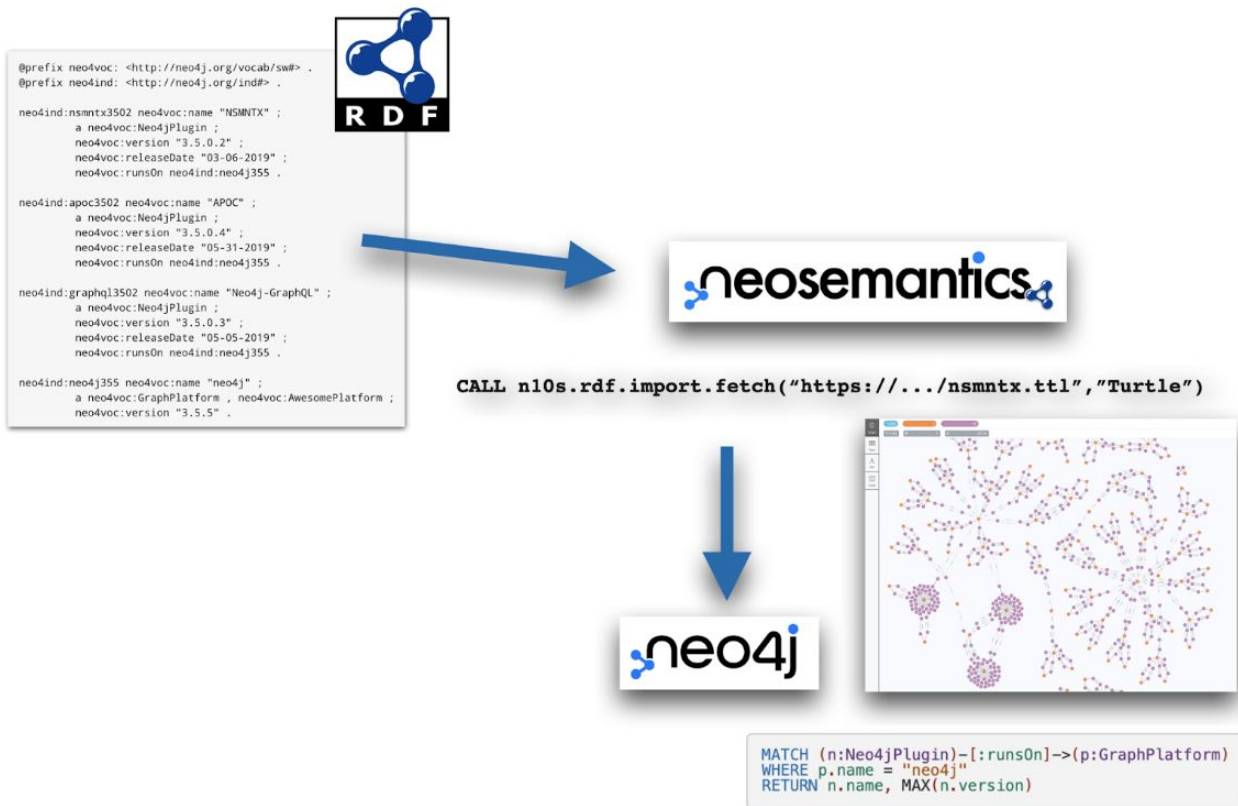


# Going Meta #12: Importing RDF data into Aura with Python + RDFLib

# What we know... (n10s)



# The “limitation”

```
Cypher Copy to Clipboard Run in Neo4j Browser
```

```
CALL n10s.graphconfig.set({ handleVocabUri: "IGNORE" });
```

```
Cypher Copy to Clipboard Run in Neo4j Browser
```

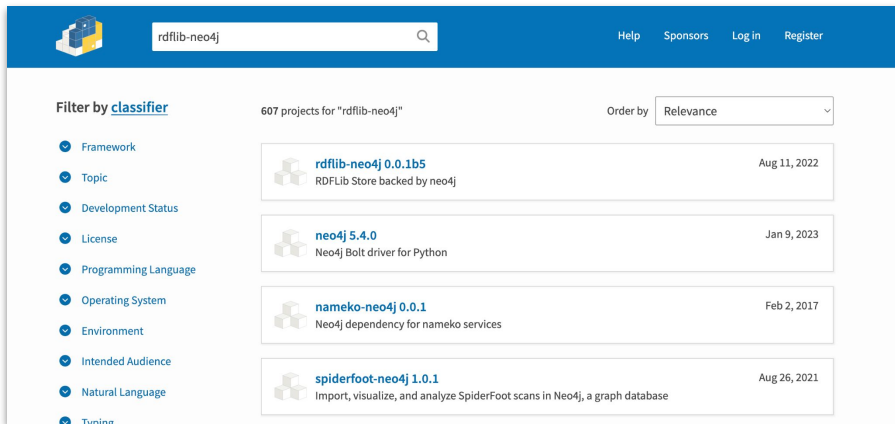
```
CALL n10s.rdf.import.fetch("https://github.com/neo4j-labs/neosemantics/raw/3.5/docs/rdf/nsmntx.ttl", "Turtle");
```



n10s deployed on the server side

# The solution: rdflib-neo4j package

<https://pypi.org/search/?q=rdflib-neo4j>



The screenshot shows the PyPI search results for 'rdflib-neo4j'. The search bar at the top contains 'rdflib-neo4j'. Below the search bar, there are filters on the left and a list of results on the right. The results list shows four packages:

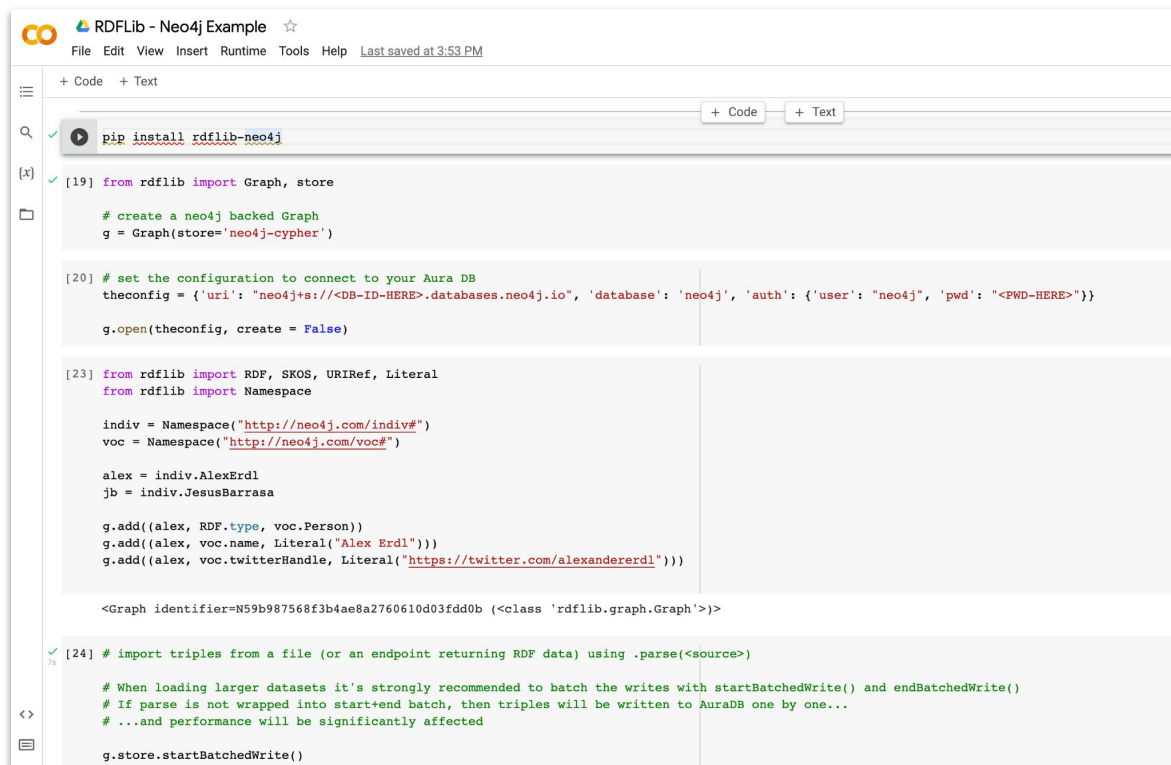
Package Name	Description	Release Date
rdflib-neo4j 0.0.1b5	RDLib Store backed by neo4j	Aug 11, 2022
neo4j 5.4.0	Neo4j Bolt driver for Python	Jan 9, 2023
nameko-neo4j 0.0.1	Neo4j dependency for nameko services	Feb 2, 2017
spiderfoot-neo4j 1.0.1	Import, visualize, and analyze SpiderFoot scans in Neo4j, a graph database	Aug 26, 2021

```
pip install rdflib-neo4j
```

```
from rdflib import Graph, store

# create a neo4j backed Graph
g = Graph(store='neo4j-cypher')
```

# Let's see it in action



The screenshot shows the Neo4j Neo4j Example application interface. The top bar includes the Neo4j logo, the title "RDFlib - Neo4j Example", and a star icon. Below the title is a menu bar with "File", "Edit", "View", "Insert", "Runtime", "Tools", and "Help". The main area is a Jupyter Notebook with the following code:

```
pip install rdflib-neo4j

[19] from rdflib import Graph, store

# create a neo4j backed Graph
g = Graph(store='neo4j-cypher')

[20] # set the configuration to connect to your Aura DB
theconfig = {'uri': "neo4j+s://<DB-ID-HERE>.databases.neo4j.io", 'database': 'neo4j', 'auth': {'user': "neo4j", 'pwd': "<PWD-HERE>"}}
g.open(theconfig, create = False)

[23] from rdflib import RDF, SKOS, URIRef, Literal
from rdflib import Namespace

indiv = Namespace("http://neo4j.com/indiv#")
voc = Namespace("http://neo4j.com/voc#")

alex = indiv.AlexErdl
jb = indiv.JesusBarrasa

g.add((alex, RDF.type, voc.Person))
g.add((alex, voc.name, Literal("Alex Erdl")))
g.add((alex, voc.twitterHandle, Literal("https://twitter.com/alexandererdl")))

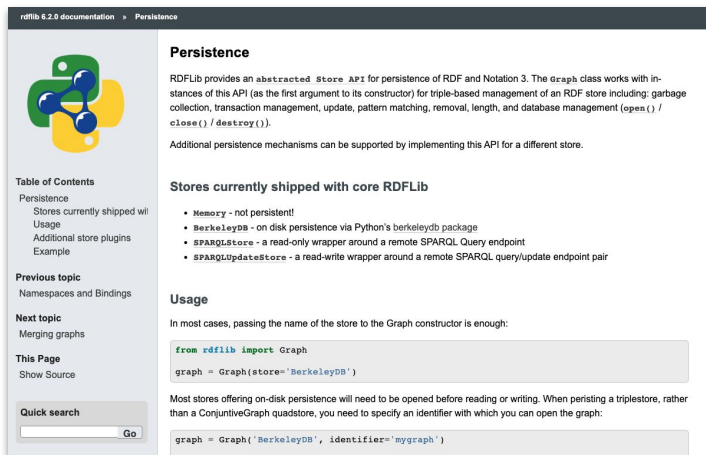
<Graph identifier=N59b987568f3b4ae8a2760610d03fdd0b (<class 'rdflib.graph.Graph'>)>

[24] # import triples from a file (or an endpoint returning RDF data) using .parse(<source>)

# When loading larger datasets it's strongly recommended to batch the writes with startBatchedWrite() and endBatchedWrite()
# If parse is not wrapped into start+end batch, then triples will be written to AuraDB one by one...
# ...and performance will be significantly affected

g.store.startBatchedWrite()
```

# Under the hood



The screenshot shows the 'Persistence' section of the rdflib 6.2.0 documentation. It includes a sidebar with a 'Table of Contents' and a 'Quick search' bar. The main content area is titled 'Persistence' and describes the abstracted Store API. It lists 'Stores currently shipped with core RDFLib' and provides a 'Usage' section with code examples for initializing a Graph with a BerkeleyDB store.

**Persistence**

RDFLib provides an abstracted `Store` API for persistence of RDF and Notation 3. The `Graph` class works with instances of this API (as the first argument to its constructor) for triple-based management of an RDF store including: garbage collection, transaction management, update, pattern matching, removal, length, and database management (`open()` / `close()` / `destroy()`).

Additional persistence mechanisms can be supported by implementing this API for a different store.

**Stores currently shipped with core RDFLib**

- `Memory` - not persistent!
- `BerkeleyDB` - on disk persistence via Python's `berkeleydb` package
- `SPARQLStore` - a read-only wrapper around a remote SPARQL Query endpoint
- `SPARQLUpdateStore` - a read-write wrapper around a remote SPARQL query/update endpoint pair

**Usage**

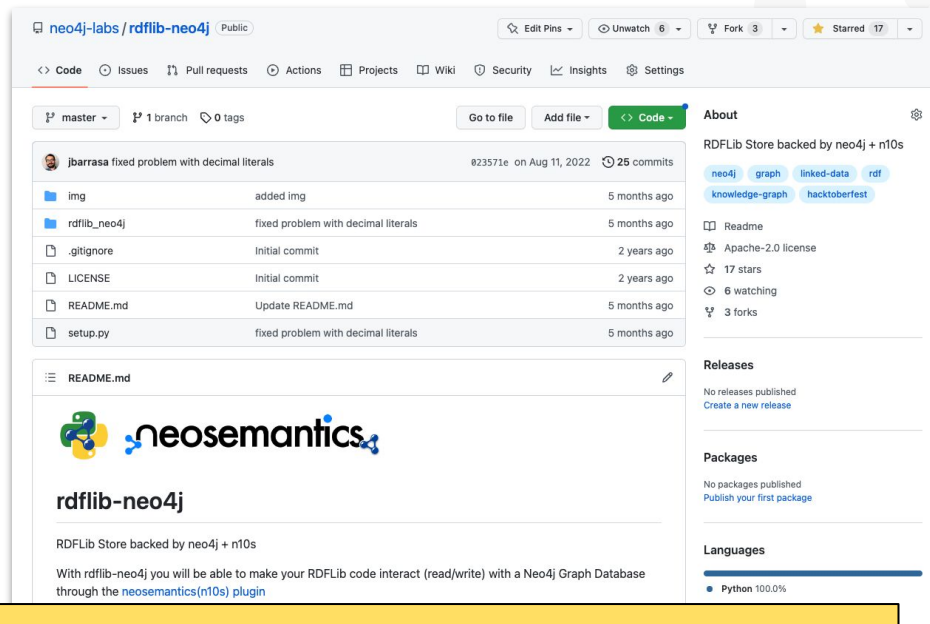
In most cases, passing the name of the store to the `Graph` constructor is enough:

```
from rdflib import Graph
graph = Graph(store='BerkeleyDB')
```

Most stores offering on-disk persistence will need to be opened before reading or writing. When persisting a triplestore, rather than a `ConjunctiveGraph` quadstore, you need to specify an identifier with which you can open the graph:

```
graph = Graph('BerkeleyDB', identifier='mygraph')
```

<https://rdflib.readthedocs.io/en/stable/persistence.html>



The screenshot shows the GitHub repository for neo4j-labs/rdflib-neo4j. It displays the repository name, a list of files (img, rdflib\_neo4j, .gitignore, LICENSE, README.md, setup.py), and a commit history table. The README section is visible, showing the neosemanantics logo and the project description: 'RDFLib Store backed by neo4j + n10s'. It also includes a 'Releases' section with 'No releases published' and a 'Packages' section with 'No packages published'.

neo4j-labs / rdflib-neo4j Public

<> Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 1 branch 0 tags

Go to file Add file <> Code About

RDLib Store backed by neo4j + n10s

neo4j graph linked-data rdf knowledge-graph hacktoberfest

Readme Apache-2.0 license 17 stars 6 watching 3 forks

Releases No releases published Create a new release

Packages No packages published Publish your first package

Languages Python 100.0%

File	Commit	Time
img	added img	5 months ago
rdflib_neo4j	fixed problem with decimal literals	5 months ago
.gitignore	Initial commit	2 years ago
LICENSE	Initial commit	2 years ago
README.md	Update README.md	5 months ago
setup.py	fixed problem with decimal literals	5 months ago

RDLib Store backed by neo4j + n10s

With rdflib-neo4j you will be able to make your RDFLib code interact (read/write) with a Neo4j Graph Database through the `neosemanantics(n10s)` plugin

Shall we have a look at the source?